A spatio-temporal entropy-based approach for the analysis of cyber attacks (demo paper)

Thibaud Mérien
Xavier Bellekens
David Brosset
Christophe Claramunt

© **2018** Copyright **held by the owner/author(s).**This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in SIGSPATIAL '18 Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems pp. 564-567

http://dx.doi.org/10.1145/3274895.3274921}.

A Spatio-temporal Entropy-based Approach for the Analysis of Cyber Attacks (Demo Paper)

Thibaud Mérien
Naval Academy Research Institute
Lanveoc, France
thibaud.merien@ecole-navale.fr

Xavier Bellekens
Division of Cyber Security, Abertay University
Dundee, United Kingdom
x.Bellekens@abertay.ac.uk

David Brosset
Naval Academy Research Institute
Lanveoc, France
david.brosset@ecole-navale.fr

Christophe Claramunt
Naval Academy Research Institute
Lanveoc, France
christophe.claramunt@ecole-navale.fr

ABSTRACT

Computer networks are ubiquitous systems growing exponentially with a predicted 50 billion devices connected by 2050. This dramatically increases the potential attack surface of Internet networks. A key issue in cyber defense is to detect, categorize and identify these attacks, the way they are propagated and their potential impacts on the systems affected. The research presented in this paper models cyber attacks at large by considering the Internet as a complex system in which attacks are propagated over a network. We model an attack as a path from a source to a target, and where each attack is categorized according to its intention. We setup an experimental testbed with the concept of honeypot that evaluates the spatio-temporal distribution of these Internet attacks. The preliminary results show a series of patterns in space and time that illustrate the potential of the approach, and how cyber attacks can be categorized according to the concept and measure of entropy.

CCS CONCEPTS

• **Security and privacy → Network security**; **Network security**; • **Networks → Network monitoring**;

KEYWORDS

Cyber attacks, Entropy, Spatial analysis

1 INTRODUCTION

Over the Internet all connected systems are very likely to be regularly attacked by thousands of malicious nodes [6]. This is especially the case for newly connected systems with the purpose of testing their capacity to resist to a given attack. In order to develop appropriate protection measures, a key issue is to understand as much as possible the attackers' intentions, and possibly to identify the perpetrator. So far different models have been set up and implemented to analyze the behavior of Internet networks [7, 15]. Several related works provide visual and global representations of attack flows on a large scale [2, 13, 16].

In order to breakthrough the obvious willingness of the attackers to remain anonymous, one solution is to compromise the attacker by pushing him to perform some attacks on some predefined, selected targets. Networks of compromised computers are called botnets and usually generate a wide panel of cyber-attacks such as Distributed Denial of Service or fishing [4, 9, 14]. Several methods have been designed and implemented to detect and analyze bots by either evaluating the network behaviour [12] or using clustering techniques [5]. The research developed in this paper provides a different approach. By setting up a series of honeypots whose objective is to act as new Internet nodes that generate a flow of attacks, our aim is to study the spatio-temporal and semantic distributions of these attacks. The Internet is modelled as a large graph, attacks are categorised and a series of entropy measures of the distribution of these attacks are computationally derived. The patterns that emerge in space and time show a series of trends that illustrate the potential of the approach. These provide several valuable insights on the possible categories, origins and locations of these cyber-attacks. The rest of this paper is organized as follows. Section 2 introduces the main principles of our modelling approach and the role of the entropy measures. Section 3 develops the case study and describes the setup of the different honeypots. Section 4 concludes the paper and outlines further work.

2 MODELLING APPROACH

The first objective of this research is to model part of the Internet activity and analyse abnormal behaviours identified as attacks and their associated intentions. Let us consider an attacker (h_{atk}) who wants to disrupt the current operations of a given system and performs an aggressive action on the targeted system (S_{trg}) property of a human target (h_{trg}). The attacker is bound to the target with the intention of his action as shown in Figure 1.

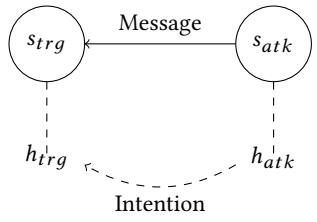

Figure 1: Network model - Intentions

In the context of the Internet, numerous systems are likely to play an implicit role in the path followed by an attack to reach its target. However, the path taken by the packets is often impossible to retrace. This leads us to introduce the concept of a black box, that

models these unknown systems between the attacker and target and as it is also almost impossible to know the exact number of systems involved. Figure 2 shows the different components of an Internet network from the attacker to his target.

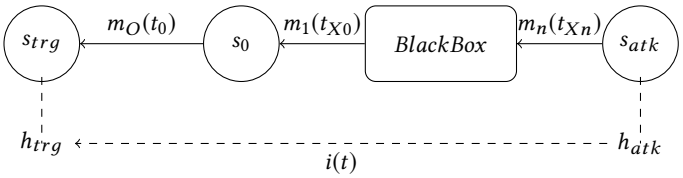

Figure 2: Network model - Attacker & Target

More formally let S be the set of systems and M the set of network messages. The notion of graph formally defines a network path between the attacker system (s_{atk}) and his targeted system (s_{trg}). The path at a time t composed by the systems $\in S$ used by the attacker (h_{atk}) to transmit a message $\in M$ to the victim h_{trg} is formally defined by the Equation 1, $G_{atk}(t) = (S, M)$.

$$G_{trg,atk}(t) = s_{trg}, m_0(t), s_0, BB[t_0, t_{n-1}], m_n(t_n), s_{atk} \quad (1)$$

Where $s_0 \in S$ the system is connected to the victim system $s_{trg} \in S$ and messages between systems are represented by $m(t) \in M$. The set of unknown systems and messages are modeled by the black box BB.

Since the attacking system and the systems composing the black box are unknown, the attack path $G_{trg,atk}(t)$ is defined as follows.

$$G_{trg,atk}(t) = s_{trg}, m_0(t), s_0^{atk} \quad (2)$$

Nodes are defined by systems and edges as messages, *i.e.*, network packets. Several useful information can be extracted from these nodes and edges. For instance, using the IP address of the nodes can help to extract their geolocation. The original intention of the attacker can also be derived from the messages extracted and categorised accordingly.

Let us define the following functions $ipv4$, $country(s)$, $city(s)$, $coord(s)$, $s24(s)$, $s16(s)$ returning, the IP address, country, city, coordinates, /24 subnet and the /16 subnet of a given system $s \in S$ respectively. The function $blacklisted(s)$ returns 'true' if an IP address is included in a known list of IP addresses that are considered suspicious, and 'false' otherwise. The function $intent(m)$ returns the intention of the attacker using the content of the message $m \in M$. From the principles represented by this modelling approach, the next step is to analyse the activity and traffic of a given Internet system, the targets and possible intentions used for anomaly detection.

The notion of entropy introduced in the seminal work of Shannon [10] denotes how much choice is involved in the selection of a statement, that is, the measure of information diversity or entropy H defined by Equation 3.

$$H_c = -K \sum p_i \log p_i \quad (3)$$

where K is a positive constant and $p_i = \frac{N_i}{N}$ is the ratio of the total number N_i of entities of the class i over the total number N of entities.

We applied the measure of entropy to the different attack intentions and categories identified. Although the notion of entropy has been already extended to the spatial dimension [1, 3] the initial measure of entropy has been here considered as the distribution of the diversity of attacks as they will appear in space over a given period of time. In order to do so, network messages have been classified into six different intentions : network infrastructure, DNS (Domain Name Server), ICS (Industrial Control Systems), Web, Control and File Sharing. Table 1 summarizes these intentions in function of the protocol.

Table 1: Protocol categorization per intention

	Intention	Protocols
Information gathering	Network Infrastructure	ICMP ; SIP ; SNMP ; SSDP
	DNS	DNS ; LLMNR ; MDNS ; NBNS
	ICS	BACnet ; DNP3.0 ; IPMI ; XDMCP ; XTACACS
	Web	HTTP
Attacks	Control	SSH ; SSHv2
	File Sharing	FTP ; TFTP

The entropy then characterizes network message intentions. In other words, the measure of entropy allows us to characterize a given geolocated IP address behavior. An IP address can be associated with a continent, country or city, this allows us to characterize the cyber behavior of a given geographical area. The measure of entropy as introduced in Equation 4 then evaluates the diversity of attacks received by a given target and IP address as follows:

$$H_c = -K \sum p_{intention} \log p_{intention} \quad (4)$$

3 IMPLEMENTATION

We applied the mechanism of honeypot, that is, a security tool that can be analysed, probed, attacked and compromised without risk for the network infrastructure [8, 11]. Honeypots are often used to deceive attackers, study attacking methods as well as obtain new and current malware samples. There are different varieties of honeypot, which can be grouped into three categories: low interaction honeypots, medium interaction honeypots and high interaction

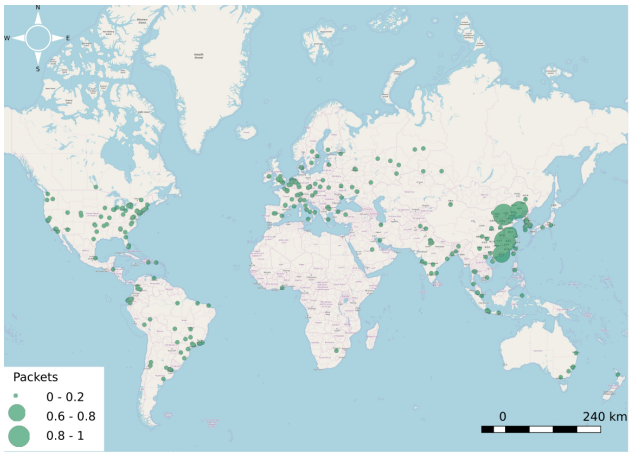

Figure 3: Number of packets for all IPs : London Honeypot

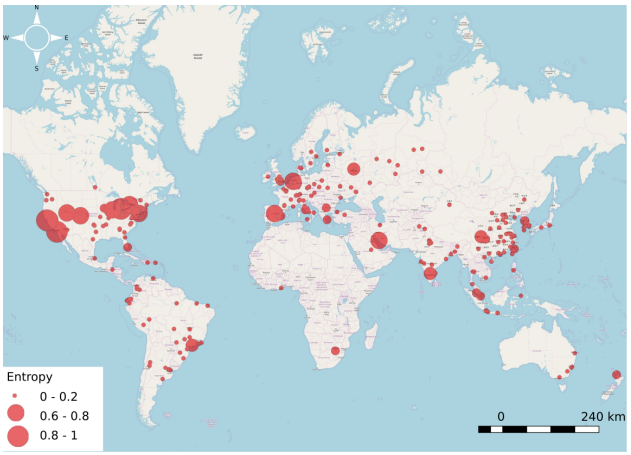

Figure 4: Entropy for all IPs : London Honeypot

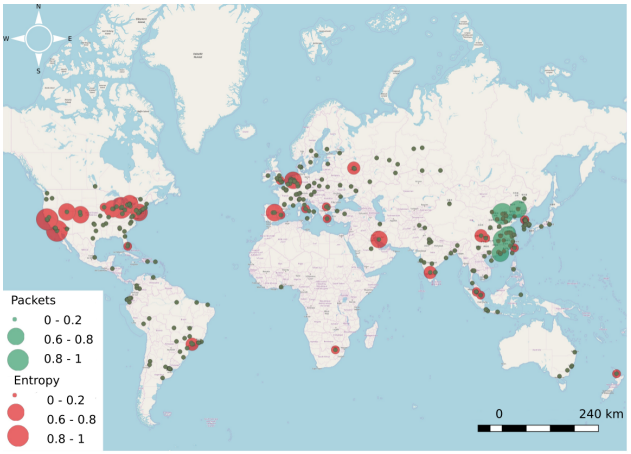

Figure 5: Entropy and number of packet for all IPs : London Honeypot

honeypots. Low interaction honeypots have limited interactions between the attackers and the system. Services are only emulated. Low interaction honeypots present the advantage of a low level of risk due to limited interactions. Medium interaction honeypots are deprived of an operating system, they emulate complex services enabling user interactions. Finally, high interaction honeypots are the most complex type of honeypots. These run a full operating system including services and a complex configuration. The main advantage of high interaction honeypots is that services are not emulated, hence, attackers are interacting with a real target while the owners can track the attackers by capturing all their interactions.

3.1 Honeypot

For the experimentation, a high interaction honeynet was set up to gather as much information as possible. In order to obtain consistent and identical data, 5 virtual private servers were purchased in 5 different locations (Fremont, Newark, London, Tokyo and Singapore). In order to correlate data from multiple sources, each honeypot was deployed with the same configuration. They run an Ubuntu 16.04 LTS with an SSH server (openSSH), an FTP server (Pure-ftpd) and a web server (Apache2). A daily PCAP file stores all network messages and 10 Gb of raw data was generated within a month. From incoming raw data, respective locations were derived using the GeoIP API based on the GeoLite database. Intentions are attributed according to the protocol of the package. In the context of this experiment, network messages have been classified into six intention categories introduced in the previous section.

3.2 Data analysis

The entropy of the intentions as well as the number of packets have been derived per city. A Python script executes SQL queries on the database containing the honeypot data, then performs the entropy calculation and visualises the outputs using the free GIS software QGIS. The data analysed and presented in this section came from the HoneyPot located in London between August 7 to August 13, 2017. Figure 3 shows the packets volume per city. A relative continuity of attack flows has been observed over the given period, throughout the whole week, day and night. Three noteworthy groups emerge: East Asia, the US and Western Europe. Hotspots are clearly apparent in large cities and reflect machine concentrations over the Internet in these cities. East Asia stands as the prominent region, with a higher number of incoming nodes and the biggest packet volumes.

Another trend that appears from East Asia is an important proportion of control attacks, this being a trend often observed in related studies. In Figure 4, each spot represents the entropy of the intentions per town. Overall, the larger the spot, the larger the diversity of attack intentions. It appears that the diversity of attacks is relatively large in the US and in some specific locations in Europe, while it is relatively low in East Asia this reflecting a concentration of attack intentions. Figure 5 combines the data presented in Figure 3 and Figure 4. The data is normalized and then projected on the map. This figure illustrates areas where the diversity of intentions is low, where the number of packets is high and conversely. The main trend that appears in Figure 5 has been observed on all the

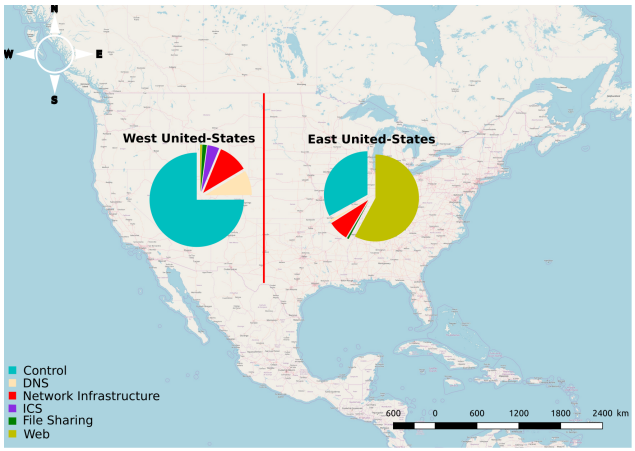

Figure 6: Zoom In : US attacks intention

honeypots implemented. As seen previously, some similar patterns emerge in terms of concentration and diversity of intention attacks and packet volumes.

At a regional scale, additional analyses have been performed. For instance, Figure 6 illustrates the trends that appear in the US, West and East coast. One can observe that the distribution of intentions is totally different although the criterion of high entropy and low packet quantity is the same. For instance, control-based attacks are prominent in the West while Web-based attacks are the main attack category that emerges from the East. It also appears that attacks are overall much more aggressive from the West Coast (i.e., control), so this is a pattern that deserves further exploration.

These preliminary patterns that emerge show that the mechanism of honeypot is a valuable solution to observe the distribution of these attack patterns, and this at different scales of observation. Over the given period anomalies have not been observed but the framework might also deliver some valuable patterns over longer periods of time.

4 CONCLUSION

The significant and continuous growth of the number of devices connected to the Internet considerably increases the risk of cyber-attacks. While several projects and frameworks have been oriented to the analysis of cyber attacks at either the local or global levels there is a need for the development of operational frameworks to analyse and identify the origins of cyber attacks on the Internet. The preliminary and experimental research presented in this paper introduces a computational approach whose objective is to qualify and understand the distribution of cyber attacks on the Internet. Several honeypots have been implemented and act as fake clients that generate massive attacks. Incoming attacks are categorised in intentions and their potential locations are approximated. A series of diversity measures have been implemented. The whole framework characterizes the semantics and diversity of these attacks in space and time. The potential of the approach is illustrated by a series of experimental patterns that emerge at the global and regional scales. Amongst many directions still to explore we plan to observe to which degree some sequences of attacks might be observed or not for a given target IP, and if this kind of regularity is replicated over different target IPs. We also plan to conduct further studies and experiments to analyse the different attack categories identified as well as apply machine learning technique to explore possible attack routes as well as unusual pattern.

REFERENCES

[1] M. Batty. Spatial entropy. *Geographical analysis*, 6:1–31, 1974.
[2] D. Brosset, C. Cavelier, Costé B., Y. Kermarrec, J. Lartigaud, and P. Laso. Cr@ck3n: A cyber alerts visualization object. In *Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA), 2017 International Conference On Cyber Situational Awareness, Data Analytics And Assessment*, pages 1–2. IEEE, 2017.
[3] C. Claramunt. A spatial form of diversity. In *International Conference on Spatial Information Theory*, pages 218–231. Springer LNCS 3693, 2005.
[4] A. Cook, A. Nicholson, H. Janicke, Maglaras L., and R. Smith. Attribution of cyber attacks on industrial control systems. *EAI transactions Industrial Networks and Intelligent Systems*, 3(7):151–158.
[5] G. Gu, R. Perdisci, J. Zhang, and Lee W. Botminer: Clustering analysis of network traffic for protocol-and structure-independent botnet detection. In *USENIX security symposium*, volume 5, pages 139–154, 2008.
[6] E. Hodo, X. Bellekens, A. Hamilton, C. Tachtatzis, and R. Atkinson. Shallow and deep networks intrusion detection system: A taxonomy and survey. *arXiv preprint arXiv:1701.02145*, 2017.
[7] B. Jasiul, M. Szpyrka, and J. Śliwa. Detection and modeling of cyber attacks with petri nets. *Entropy*, 16(12):6602–6623, 2014.
[8] I. Mokube and M. Adams. Honeypots: concepts, approaches, and challenges. In *Proceedings of the 45th annual southeast regional conference*, pages 321–326. ACM, 2007.
[9] A. Nicholson, T. Watson, P. Norris, Duffy A., and R. Isbell. A taxonomy of technical attribution techniques for cyber attacks. In *European Conference on Information Warfare and Security*, page 188. Academic Conferences International Limited, 2012.
[10] C. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.
[11] L. Spitzner. *Honeypots: Tracking Hackers*, volume 1. Addison-Wesley Reading, 2003.
[12] W. Strayer, D. Lapsely, Walsh R., and C. Livadas. Botnet detection based on network behavior. In *Botnet Detection*, pages 1–24. Springer, 2008.
[13] M. Syamkumar, Durairajan R., and P. Barford. Bigfoot: A geo-based visualization methodology for detecting bgp threats. In *Visualization for Cyber Security (VizSec), IEEE Symposium on*, pages 1–8. IEEE, 2016.
[14] O. Thonnard, Mees W., and M. Dacier. On a multicriteria clustering approach for attack attribution. *ACM SIGKDD Explorations Newsletter*, 12(1):11–20, 2010.
[15] J. Tölle and Niggemann O. Supporting intrusion detection by graph clustering and graph drawing. In *Proceedings of Third International Workshop on Recent Advances in Intrusion Detection RAID 2000*, 2000.
[16] M Withall, Phillips I., and Parish D. Network visualisation: a review. *IET communications*, 1(3):365–372, 2007.